

**Adv Pharm Bull**. 2025;15(2):467-473 doi: 10.34172/apb.025.43852 https://apb.tbzmed.ac.ir



### **Short Communication**



# Assessing the Potential of Generative Artificial Intelligence Models to Assist Experts in the Development of Pharmacokinetic Models

Sergio Sánchez-Herrero<sup>1</sup>, Laura Calvet Liñan<sup>2</sup>

<sup>1</sup>Department of Computer Science, Multimedia and Telecommunication, Universitat Oberta de Catalunya, Barcelona, 08018, Spain <sup>2</sup>Department of Telecommunications & Systems Engineering, Universitat Autònoma de Barcelona, Sabadell, 08202, Spain

#### Article info

#### Article History:

Received: October 7, 2024 Revised: May 31, 2025 Accepted: June 1, 2025 epublished: June 3, 2025

### **Keywords:**

Artificial intelligence, Generative models, Pharmacokinetics, ChatGPT, Modelling

### **Abstract**

**Purpose:** This study explores the potential of generative AI models to aid experts in developing scripts for pharmacokinetic (PK) models, with a focus on constructing a two-compartment population PK model using data from Hosseini et al.

*Methods:* Generative AI tools ChatGPT v3.5, Gemini v2.0 Flash and Microsoft Copilot free could help PK professionals— even those without programming experience—learn the programming languages and skills needed for PK modeling. To evaluate these free AI tools, PK models were created in R Studio, covering key tasks in pharmacometrics and clinical pharmacology, including model descriptions, input requirements, results, and code generation, with a focus on reproducibility.

**Results:** ChatGPT demonstrated superior performance compared to Copilot and Gemini, highlighting strong foundational knowledge, advanced concepts, and practical skills, including PK code structure and syntax. Validation indicated high accuracy in estimated and simulated plots, with minimal differences in clearance (Cl) and volume of distribution (V c and V p) compared to reference values. The metrics showed absolute fractional error (AFE), absolute average fractional error (AAFE), and mean percentage error (MPE) values of 0.99, 1.14, and -1.85, respectively.

**Conclusion:** These results show that generative AI can effectively extract PK data from literature, build population PK models in R, and create interactive Shiny apps for visualization, with expert support.

# Introduction

Generative artificial intelligence (AI) models, including ChatGPT from OpenAI, Bard from Google (Gemini), and Copilot from Microsoft, exemplify a rapidly advancing technology that has garnered significant private and public attention.<sup>1</sup> These models possess the potential to transform pharmacokinetics (PK) by offering novel research and analysis pathways. However, it is imperative to acknowledge that these technologies are still in the early stages of development and are susceptible to misuse. It is imperative to ensure meticulous integration into PK research.

PK and pharmacodynamics (PD) modelling play a crucial role in preclinical research by enabling the characterization of drug concentration-time profiles, interspecies scaling, and dose selection. Population-based approaches and mathematical models are essential tools that help pharmacologists assess drug exposure, efficacy, and safety. The use of nonlinear mixed-effects and Bayesian methods has improved the analysis of limited

clinical data, helping to predict variability, optimise dosing, and support regulatory decisions.<sup>2</sup>

Recent advancements, particularly those stemming from machine learning (ML) and generative models, have led to novel approaches for enhancing PK modelling. These approaches involve the generation of synthetic data that emulates real-world conditions and the simulation of intricate biological systems.3 When integrated with domain expertise, these models have the potential to enhance prediction accuracy, uncover latent patterns, and facilitate decision-making processes in the field of drug development.4 The utilisation of tools such as pyDarwin, which integrates ML with NONMEM, exemplifies the efficacy of hybrid methodologies.<sup>5</sup> However, its limitations - particularly in relation to complex or unfamiliar tasks - mean that it must be used with caution, validated rigorously, and applied by experts in PK and statistics alongside traditional methods to ensure reliability.6

In this context, the objective of this study is to provide a preliminary exploration of the development of generative

<sup>\*</sup>Corresponding Author: Sergio Sánchez-Herrero, Email: ssanchezherre@uoc.edu

AI, with a particular focus on examining the capabilities of ChatGPT, Copilot, and Gemini in assisting PK experts by providing simulation support in code and generating relevant knowledge for practice, learning, and research. The focus of this exploration will be on examining these AI generative models in providing simulation assistance in code and generating knowledge relevant to the practice, learning and research of experts in PK. The following two points will be analysed in this study: i) the potential of generative AI models to act as tools to assist PK experts in the execution of their professional duties; and ii) the presentation of an illustrative example by validating a population analysis developed by Hosseini et al.<sup>7</sup>

### Material and Methods

### Software

Different generative AI models were used like ChatGPT v3.5 (https://chat.openai.com; OpenAI; version as of April 4, 2024), Gemini v2.0 Flash (https://gemini.google.com; Google AI; version as of December 11, 2024), and Microsoft Copilot free (https://copilot.microsoft.com; Microsoft; version as of September 21, 2023). It is important to take into account that all generative AI models are free.

R program language (version 4.2.1) was used in the current study for data pre-processing and pop analysis implementation. The following packages were utilized: ggplot2 (version 3.5.0),<sup>8</sup> readxl (version 1.4.1),<sup>9</sup> Shiny (version 1.7.2)<sup>10</sup> and deSolve (version 1.40).<sup>11</sup> Unless otherwise specified, default parameters were used for each programming function.

AI generative models were employed to provide descriptions and responses to queries according to the

workflow outlined in Figure 1. Utilizing the generated R code, subsequent inquiries were made to enhance the functionalities of the PK model.

# Pharmacokinetics model for simulation

To evaluate the use of generative AI in pharmacokinetic modelling, we replicated the two-compartment model from Hosseini et al<sup>7</sup> (Case 1) using R; dataset and details are available in the original Supplementary material.

# Data fitting re-estimation

PK parameters were estimated using the mapbayr package in R (v0.10.0), which applies a maximum a posteriori Bayesian method to models built with mrgsolve (version 1.4.1). Initial values are listed in Table 1.

### **Model Validation**

The resulting code was executed to predict twocompartment model parameters, which were then compared to reference values from Hosseini et al.<sup>7</sup> For reproducibility and verification, prediction performance was assessed using standard metrics: prediction error (PE), mean percentage error (MPE), average-fold error

Table 1. Initial conditions for the two-compartment model

Variable	Estimated	Initial	Units	Min	Max
V <sub>1</sub>	True	40	mm/kg	10	200
$V_2$	True	40	mm/kg	10	200
C <sub>I</sub>	True	5	mm/day/kg	1	30
Q	True	10	mm/day/kg	1	100
V <sub>m</sub>	False	0	μg/day/kg	0	120
K <sub>m</sub>	False	5	μg/mm	0.1	100

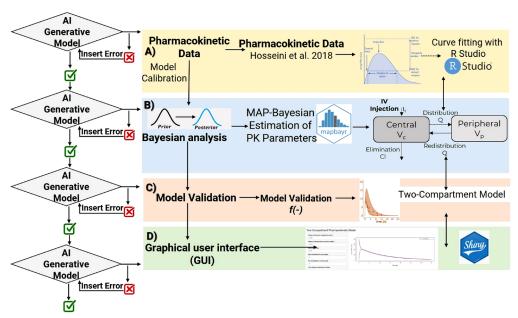


Figure 1. Workflow followed to build a generative AI application for developing a two-compartment model. AI generative model developed and designed to process and generate text in response to user queries and requests. A) Pharmacokinetic data by Hosseini et al.<sup>7</sup> B) Two-Compartment model was calibrated with experimental datasets and was optimized and parameter uncertainty and variability were characterized within a MAP-Bayesian estimation of PK parameters framework (mapbayr). The final model was subjected to evaluation/validation. D) The final PK population model was converted to a web-based graphical user interface with Shiny R package<sup>10</sup>

(AFE), and absolute average-fold error (AAFE). Predictive performance was considered satisfactory if AFE and AAFE values fell within the 0.8–1.25-fold range.<sup>13</sup>

A visual predictive check (VPC) was conducted to assess the predictive performance of the PK model. Observed concentrations were dose-normalized to a standard dose using a Monte Carlo simulation approach implemented in R, incorporating generative models along with the parameter estimates and variability described in Table 2.

Additionally, goodness-of-fit (GOF) plots were performance to assess model adequacy. R can generate a variety of GOF plots, including graphs like observations versus individual and population model predictions or ETA distribution from estimation results.

### GitHub repository

All steps, data analysis procedures, results, and accompanying analyses presented in this study are made openly accessible through an open repository for the sake of transparency, reproducibility, and community engagement. The URL for the guide is: https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/tree/main.

#### Results

### Two-compartment model using generative model

Various generative AI tools (Copilot, Gemini and ChatGPT) were prompted to generate a two-compartment model based on Hosseini et al.<sup>7</sup> Outputs were compared to published simulations and real data. All prompts, steps and code are available at https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/tree/main.

Each generative AI model produced a slightly different R code for developing the PK model, leading to variations in the results. We encountered several errors while executing the generated R codes, which were resolved by sharing the issues with the respective AI models. Gemini and Microsoft Copilot did not produce successful results and they were excluded from further steps. It is important to note that specificity in input generally yields better results. However, discrepancies across AI models were mainly related to the graph legends. The text input used to obtain

**Table 2.** Common values and variability of model parameters for the simulated data

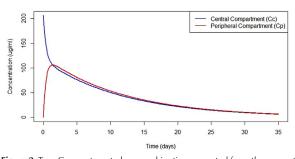
	Simulated data two-compartment model (Pls 5th and 95 <sup>th</sup> )	
Dose	10 mg/100 mg	
Population	1000 subjects	
Population parameters	V c=fit l V p=fit l CLt=fit l/h CLd=fit l/h	
Inter-subject variability	variability population result	
Residual error	$V ar = (0.01 + 0.1 \cdot IPRED)^2$	

Figure 2 from the PK model is available at https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/blob/main/generative\_AI\_text/model.txt

The same code and asking text were adapted by AI generative models using the Shiny package to create a user interface. The additional text provided to the AI models was: "The code you generate, I want it in Shiny, with inputs for the parameter definition and the time". R code to obtain Shiny application is located in https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/blob/main/R/shinyall.R.

Parameter estimation for population analysis was most straightforward using mapbayr, as generative AI models readily interpreted its Bayesian framework.<sup>12</sup> The results of the pooled fitting are illustrated in Figure 3 for 10 mg and 100 mg, and the parameter estimates for both pooled, Figure 4 for individual mapbayr predicted curves and the group specific fittings are listed in Table 3, demonstrating relatively low standard errors. The steps followed to obtain the results are available at https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/blob/main/generative\_AI\_text/pop.txt.

On the other hand, the model and parameter estimates were used together to project PK for an alternate dosing regimen and explore potential PK variability. The parameter estimates from the two-compartment model were saved to use them for projecting the PK profile of a 10 mg/kg 4qw (weekly dosing for 4 weeks) IV dosing regimen under different conditions, listed in Table 4. The population simulation based on variably exposed in Table 4 shows AUC (0-28) similar to the reference too. Figure 5 provides a summary of the projective simulations. Prompts text could be located in https://github.com/sersanchezherrero/Generative AI PKpop Model/blob/



**Figure 2.** Two-Compartment pharmacokinetics generated from the generative AI code located in https://github.com/sersanchezherrero/Generative\_AI\_ PKpop\_Model/blob/main/R/two\_compartment\_model.R. Blue line central compartment concentration. Red line peripheral compartment concentration. The two-compartment model follows the expected pharmacokinetic behaviour, demonstrating proper distribution and elimination phases

Table 3. Fitting combined pooled results Hosseini et al<sup>7</sup> versus PhysPK

Name	Initial	Fit-group	AI generative Model	Units
V1	40	49.15+/-3.8	49.86+/-1.43	ml/kg
V2	40	34.61 +/-5.2	33.14+/-1.33	ml/kg
CL	5	6.89+/-0.2	7.02+/-0.78	ml/day/kg
Cld	10	45.5+/-17.4	50.55 +/-0.5	ml/day/kg

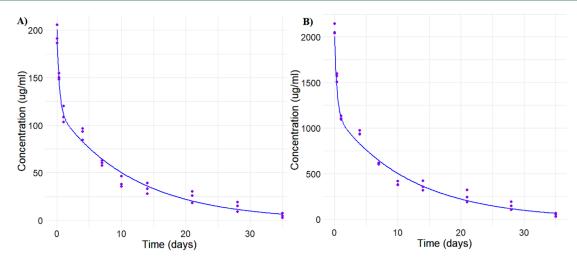


Figure 3. Pooled fit graphs generated with from generative AI model code generated by dose according to intravenous administration, (A) 10 mg. and (B) 100 mg. Blue: Simulated fit data. Purple: Real data. Concentration-time profiles (A-B) demonstrate consistent pharmacokinetic trends across observed times, with smooth curves indicating stable model behaviour and reliable predictions

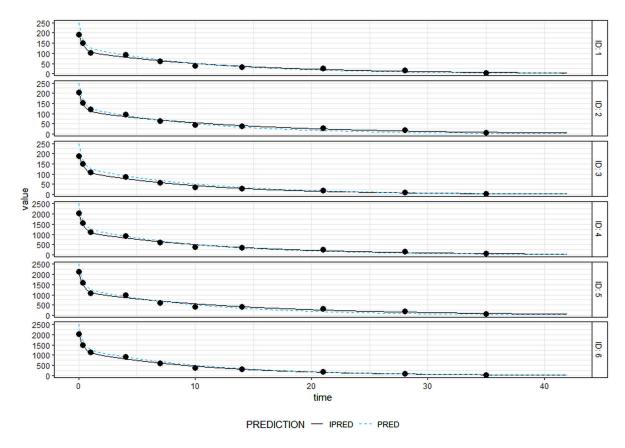


Figure 4. Individual concentrations and individual predicted curves for the pharmacokinetics models generated by mapbayr package. Black dots are real data. Black lines are individual predictions (IPRED). Dash-blue lines are population predictions (PRED). On the right of the figure are located ID patient. The consistent overlap of trajectories across subjects indicates reproducible pharmacokinetic behaviour and model reliability

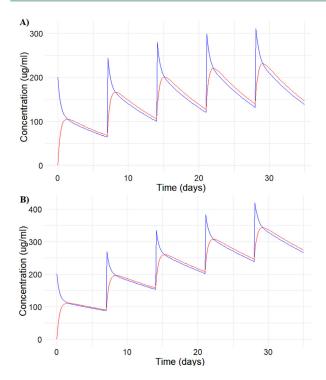
**Table 4.** Fitting combined pooled results Hosseini et al<sup>7</sup> versus PhysPK

	Scenario	Ref. AUC (0-28)	AI AUC (0-28)
Simulation	CL1 = 7.02 (Fit)	4162.31	4068.77
Simulation	CL2 = 0.5 x Cl1 = 3.51	5759.83	5652.29

main/generative\_AI\_text/simulation.txt and R code in https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/blob/main/R/simu.R.

# Visual predictive check

To evaluate the effectiveness of a model's predicted performance, VPC charts were interacted with the results, with further details provided at https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/blob/main/generative\_AI\_text/vpc.txt. The graph generated by VPC in Figure 6 provides a summary of the predictive simulations.



**Figure 5.** Two-Compartment pharmacokinetic model projection for the PK profile of a 10 mg/kg 4qw (weekly dosing for 4 weeks) IV dosing regimen under A) CL1=7.02 (Fit) and B) CL2=0.5 x Cl1=3.51. Blue line: Central Compartment. Red line: Peripheral Compartment. The smooth concentration-time trajectories (A-C) reflect reliable model performance, with no aberrant patterns, indicating stable parameter estimation and predictive accuracy

### Parameter distribution profiles

The log-normal distribution of parameters holds significance as it cannot reveal the presence of subpopulations. In Figure 7, density graphs are reported. They effectively convey information regarding the variability around each parameter, affirming the suitability of the chosen distribution for this model. The generative AI models were interacted with the results, with further details provided at https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/blob/main/generative\_AI text/dis.txt

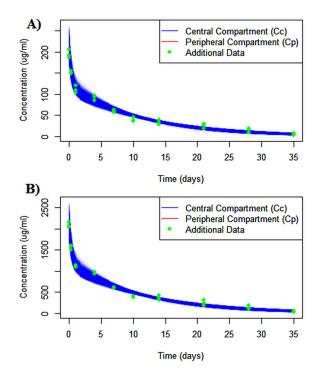
To evaluate the accuracy of model-based the AFE and AAFE values fall between 0.8 and 1.25. The precision metrics results were AFE = 0.99, AAFE = 1.14 and MPE = -1.85 for the six real patients data. The metrics showed proper accuracy based on references.<sup>13</sup>

### Goodness-of-fit

AI generative model code for developing GOF plot and ETA distribution plots are located in https://github.com/sersanchezherrero/Generative\_AI\_PKpop\_Model/blob/main/generative\_AI\_text/gof.txt. Generated plots are shown in Figure 8. Visual inspection of the precision plot (observation vs. individual and population model prediction) indicates that the predictions obtained are acceptable compared to the reference.

### Discussion

Generative AI models such as ChatGPT and Copilot are emerging as valuable tools in PK modeling. Although



**Figure 6.** Virtual predicted check (VPC) generated with R by dose according to intravenous administration. A population dataset of 1000 subjects was generated using Monte Carlo simulation method for A) 10 and B) 100 mg administration. The fitted concentration-time curves demonstrate good agreement between the model and observed data

not specifically designed for pharmacometrics, they can assist with code generation, model interpretation, and troubleshooting, supporting both experienced users and those with limited programming skills.<sup>14</sup>

Traditional PK analyses typically rely on software like NONMEM, Monolix, and WinNonlin, which offer advanced functionalities but often require specialized expertise. In contrast, generative AI tools facilitate more accessible and flexible workflows within platforms like R, helping to translate complex PK concepts into executable code and promoting interdisciplinary collaboration. 15,16

In our study, we developed a two-compartment population PK model for single-dose intravenous antibody administration using ChatGPT and Copilot. The AIgenerated models produced pharmacokinetic parameter estimates with high accuracy. For instance, clearance values (CL and CLd) were  $6.89 \pm 0.2$  and  $45.5 \pm 17.4$ , closely matching the reference values of  $7.02 \pm 0.78$  and  $50.55 \pm 0.5$ , respectively. Similarly, volumes of distribution (V1 and V2) were estimated at  $49.15 \pm 3.8$  and  $34.61 \pm 5.2$ , compared to reference values of 49.86 ± 1.43 and  $33.14 \pm 1.33$ . Model performance metrics were also within acceptable ranges, with an AFE of 0.99, AAFE of 1.14, and MPE of -1.85. These results align with previous studies, such as those by Hosseini et al., validating the ability of generative AI to accurately replicate standard population PK models.

However, these tools have important limitations. Their responses are based on pre-trained data, which may not reflect the latest advancements in pharmacometrics, and

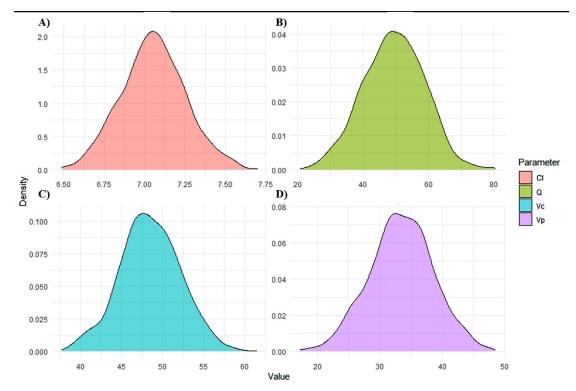
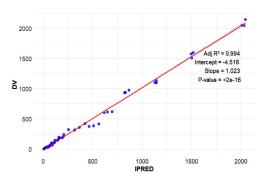


Figure 7. Density graph from Montecarlo simulation parameters. A) Clearance (CL). B) Distribution (Q). Central Volume (Vc). D) Peripheral volume (Vp). The Figure provides fixed estimates for key parameters, supporting quantitative interpretation. Such plots are essential for evaluating parameter precision in Bayesian or population modeling. The smooth, symmetric density curves (A-D) indicate well-behaved parameter distributions, suggesting reliable model estimates



**Figure 8.** Goodness-of-fit plots (PRED versus DV) for the final model. Abbreviations: DV, observed plasma concentration; iPreD, individual predicted values. Observed vs. Individual Predicted Values (IPRED) with Regression Line. The tight alignment of points along the diagonal line indicates excellent model performance, with predictions closely matching actual observations. This validation step ensures the reliability of the model for interpreting or forecasting outcomes

they lack the specialized insight of human experts. There is also a risk of generating oversimplified or incorrect outputs, particularly in complex or novel scenarios. Issues related to reproducibility, transparency, and intellectual property must be considered, and ethical use requires clear attribution and expert oversight.

Despite these challenges, ongoing advances—such as the improved reasoning and language capabilities in models like ChatGPT-4—suggest that generative AI will play an increasingly important role in PK/PD workflows. Used judiciously, these tools can streamline the early stages of model development, improve efficiency, and expand accessibility while maintaining scientific rigor through appropriate validation and human supervision. 17,18

### Conclusion

This study highlights the transformative role of AI generative models in pharmacokinetic modelling using R. These tools enhance information synthesis, collaboration, education, and code generation, making PK modeling more accessible and efficient.

By developing a two-compartment population PK model, we demonstrated that AI tools—particularly ChatGPT and Copilot—can accurately estimate key parameters (CL, V, CLd), with minimal discrepancies compared to traditional approaches. Their integration with R enables streamlined workflows, reduces manual errors, and supports reproducible research.

AI models offer significant advantages in routine model evaluation, publication preparation, and decision-making. However, they should complement—not replace—human expertise. Responsible use requires clear objectives, context, and expert oversight to ensure scientific rigor.

In summary, AI generative models represent a major advancement in pharmacometrics, with growing potential to further transform the field through continued innovation and integration.

### Acknowledgements

We would like to express our gratitude to the authors whose references were utilized in this study, which aims to juxtapose our ML methodologies with real-world data from open publications.

### **Authors' Contribution**

Conceptualization: Sergio Sánchez-Herrero and Laura Calvet Liñan.

Data curation: Sergio Sánchez-Herrero.

Formal analysis: Sergio Sánchez-Herrero and Laura Calvet Liñan.

Investigation: Sergio Sánchez-Herrero. Methodology: Sergio Sánchez-Herrero. Project administration: Laura Calvet Liñan. Resources: Sergio Sánchez-Herrero. Software: Sergio Sánchez-Herrero. Supervision: Laura Calvet Liñan. Validation: Laura Calvet Liñan.

Visualization: Sergio Sánchez-Herrero and Laura Calvet Liñan.

Writing-original draft: Sergio Sánchez-Herrero. Writing-review & editing: Laura Calvet Liñan.

# **Competing Interests**

The authors declare that they have no conflicts interests.

#### **Data Availability Statement**

All data used, generated or analyzed during this study are included in https://github.com/sersanchezherrero/Generative\_Al\_PKpop\_Model/tree/main.

### **Ethical Approval**

Not applicable.

### **Funding**

None.

### **Supplementary Files**

Supplementary File contains a zip file.

#### References

- Rossettini G, Rodeghiero L, Corradi F, Cook C, Pillastrini P, Turolla A, et al. Comparative accuracy of ChatGPT-4, Microsoft Copilot and Google Gemini in the Italian entrance test for healthcare sciences degrees: a cross-sectional study. BMC Med Educ 2024;24(1):694. doi: 10.1186/s12909-024-05630-9
- Dartois C, Lemenuel-Diot A, Laveille C, Tranchand B, Tod M, Girard P. Evaluation of uncertainty parameters estimated by different population PK software and methods. *J Pharmacokinet Pharmacodyn* 2007;34(3):289-311. doi: 10.1007/s10928-006-9046-9
- Russell SJ, Norvig P. Artificial Intelligence: A Modern Approach. 3rd ed. Pearson; 2016.
- 4. Huang X, Estau D, Liu X, Yu Y, Qin J, Li Z. Evaluating the performance of ChatGPT in clinical pharmacy: a comparative study of ChatGPT and clinical pharmacists. *Br J Clin Pharmacol* 2024;90(1):232-8. doi: 10.1111/bcp.15896
- Li X, Sale M, Nieforth K, Bigos KL, Craig J, Wang F, et al. pyDarwin: a machine learning enhanced automated nonlinear

- mixed-effect model selection toolbox. *Clin Pharmacol Ther* 2024;115(4):758-73. doi: 10.1002/cpt.3114
- Wach K, Duong CD, Ejdys J, Kazlauskaitė R, Korzynski P, Mazurek G, et al. The dark side of generative artificial intelligence: a critical analysis of controversies and risks of ChatGPT. Entrep Bus Econ Rev 2023;11(2):7-30. doi: 10.15678/eber.2023.110201
- 7. Hosseini I, Gajjala A, Bumbaca Yadav D, Sukumaran S, Ramanujan S, Paxson R, et al. gPKPDSim: a SimBiology®-based GUI application for PKPD modeling in drug development. *J Pharmacokinet Pharmacodyn* 2018;45(2):259-75. doi: 10.1007/s10928-017-9562-9
- 8. Wickham H. ggplot2. Wiley Interdiscip Rev Comput Stat 2011;3(2):180-5. doi: 10.1002/wics.147
- Wickham H, Bryan J. readxl: Read Excel Files. R Package Version 1.4.2. 2023. Available from: https://CRAN.R-project. org/package=readxl.
- Li R, Bilal U. Interactive web-based data visualization with R, plotly, and shiny (Carson Sievert). *Biometrics* 2021;77(2):776-7. doi: 10.1111/biom.13474
- Soetaert K, Petzoldt T, Setzer RW. Solving differential equations in R: package deSolve. *J Stat Softw* 2010;33(9):1-25. doi: 10.18637/jss.v033.i09
- Le Louedec F, Puisset F, Thomas F, Chatelut É, White-Koning M. Easy and reliable maximum a posteriori Bayesian estimation of pharmacokinetic parameters with the open-source R package mapbayr. CPT Pharmacometrics Syst Pharmacol 2021;10(10):1208-20. doi: 10.1002/psp4.12689
- Martínez FC, Serna J, Cuquerella-Gilabert M, Rueda-Ferreiro A, Juan AA, Calvet L. Embedding R inside the PhysPK biosimulation software for pharmacokinetics population analysis. BIO Integr 2023;4(3):97-113. doi: 10.15212/bioi-2023-0008
- Lv Z. Generative artificial intelligence in the metaverse era. Cogn Robot 2023;3:208-17. doi: 10.1016/j.cogr.2023.06.001
- Rajman I. PK/PD modelling and simulations: utility in drug development. *Drug Discov Today* 2008;13(7-8):341-6. doi: 10.1016/j.drudis.2008.01.003
- Ghayoor A, Kohan HG. Revolutionizing pharmacokinetics: the dawn of Al-powered analysis. *J Pharm Pharm Sci* 2024;27:12671. doi: 10.3389/jpps.2024.12671
- 17. Derendorf H, Meibohm B. Modeling of pharmacokinetic/ pharmacodynamic (PK/PD) relationships: concepts and perspectives. *Pharm Res* 1999;16(2):176-85. doi: 10.1023/a:1011907920641
- Massey PA, Montgomery C, Zhang AS. Comparison of ChatGPT-3.5, ChatGPT-4, and orthopaedic resident performance on orthopaedic assessment examinations. J Am Acad Orthop Surg 2023;31(23):1173-9. doi: 10.5435/ jaaos-d-23-00396